

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/853,450

DATE: 10/24/2001

TIME: 16:01:58

Input Set : A:\-24.app

Output Set: N:\CRF3\10242001\I853450.raw

3 <110> APPLICANT: Yanofsky, Martin F.
 4 Pelaz, Soraya
 5 Ditta, Gary
 6 The Regents of the University of California
 8 <120> TITLE OF INVENTION: Combinations of Genes for Producing Seed Plants
 9 Exhibiting Modulated Reproductive Development
 11 <130> FILE REFERENCE: 19452A-002400US
 13 <140> CURRENT APPLICATION NUMBER: US 09/853,450
 14 <141> CURRENT FILING DATE: 2001-05-09
 16 <160> NUMBER OF SEQ ID NOS: 61
 18 <170> SOFTWARE: PatentIn Ver. 2.1
 20 <210> SEQ ID NO: 1
 21 <211> LENGTH: 1057
 22 <212> TYPE: DNA
 23 <213> ORGANISM: Arabidopsis thaliana
 25 <220> FEATURE:
 26 <221> NAME/KEY: CDS
 27 <222> LOCATION: (124)..(894)
 28 <223> OTHER INFORMATION: APETALA1 (AP1)
 30 <400> SEQUENCE: 1
 31 ctttccaatt ggttcatacc aaagtctgag ctcttcttta tatctctctt gtagtttctt 60
 33 attggggggtc tttgttttgt ttggttcttt tagagtaaga agtttcttaa aaaaggatca 120
 35 aaa atg gga agg ggt agg gtt caa ttg aag agg ata gag aac aag atc 168
 36 Met Gly Arg Gly Arg Val Gln Leu Lys Arg Ile Glu Asn Lys Ile
 37 1 5 10 15
 39 aat aga caa gtg aca ttc tcg aaa aga aga gct ggt ctt ttg aag aaa 216
 40 Asn Arg Gln Val Thr Phe Ser Lys Arg Arg Ala Gly Leu Leu Lys Lys
 41 20 25 30
 43 gct cat gag atc tct gtt ctc tgt gat gct gaa gtt gct ctt gtt gtc 264
 44 Ala His Glu Ile Ser Val Leu Cys Asp Ala Glu Val Ala Leu Val Val
 45 35 40 45
 47 ttc tcc cat aag ggg aaa ctc ttc gaa tac tcc act gat tct tgt atg 312
 48 Phe Ser His Lys Lys Gly Lys Leu Phe Glu Tyr Ser Thr Asp Ser Cys Met
 49 50 55 60
 51 gag aag ata ctt gaa cgc tat gag agg tac tct tac gcc gaa aga cag 360
 52 Glu Lys Ile Leu Glu Arg Tyr Glu Arg Tyr Ser Tyr Ala Glu Arg Gln
 53 65 70 75
 55 ctt att gca cct gag tcc gac gtc aat aca aac tgg tcg atg gag tat 408
 56 Leu Ile Ala Pro Glu Ser Asp Val Asn Thr Asn Trp Ser Met Glu Tyr
 57 80 85 90 95
 59 aac agg ctt aag gct aag att gag ctt ttg gag aga aac cag agg cat 456
 60 Asn Arg Leu Lys Ala Lys Ile Glu Leu Leu Glu Arg Asn Gln Arg His
 61 100 105 110
 63 tat ctt ggg gaa gac ttg caa gca atg agc cct aaa gag ctt cag aat 504
 64 Tyr Leu Gly Glu Asp Leu Gln Ala Met Ser Pro Lys Glu Leu Gln Asn
 65 115 120 125
 67 ctg gag cag cag ctt gac act gct ctt aag cac atc cgc act aga aaa 552

ENTERED

RAW SEQUENCE LISTING

DATE: 10/24/2001

PATENT APPLICATION: US/09/853,450

TIME: 16:01:58

Input Set : A:\-24.app

Output Set: N:\CRF3\10242001\I853450.raw

```

68 Leu Glu Gln Gln Leu Asp Thr Ala Leu Lys His Ile Arg Thr Arg Lys
69      130      135      140
71 aac caa ctt atg tac gag tcc atc aat gag ctc caa aaa aag gag aag 600
72 Asn Gln Leu Met Tyr Glu Ser Ile Asn Glu Leu Gln Lys Lys Glu Lys
73      145      150      155
75 gcc ata cag gag caa aac agc atg ctt tct aaa cag atc aag gag agg 648
76 Ala Ile Gln Glu Gln Asn Ser Met Leu Ser Lys Gln Ile Lys Glu Arg
77 160      165      170      175
79 gaa aaa att ctt agg gct caa cag gag cag tgg gat cag cag aac caa 696
80 Glu Lys Ile Leu Arg Ala Gln Gln Glu Gln Trp Asp Gln Gln Asn Gln
81      180      185      190
83 ggc cac aat atg cct ccc cct ctg cca ccg cag cag cac caa atc cag 744
84 Gly His Asn Met Pro Pro Pro Leu Pro Pro Gln Gln His Gln Ile Gln
85      195      200      205
87 cat cct tac atg ctc tct cat cag cca tct cct ttt ctc aac atg ggt 792
88 His Pro Tyr Met Leu Ser His Gln Pro Ser Pro Phe Leu Asn Met Gly
89      210      215      220
91 ggt ctg tat caa gaa gat gat cca atg gca atg agg agg aat gat ctc 840
92 Gly Leu Tyr Gln Glu Asp Asp Pro Met Ala Met Arg Arg Asn Asp Leu
93      225      230      235
95 gaa ctg act ctt gaa ccc gtt tac aac tgc aac ctt ggc tgc ttc gcc 888
96 Glu Leu Thr Leu Glu Pro Val Tyr Asn Cys Asn Leu Gly Cys Phe Ala
97 240      245      250      255
99 gca tga agcatttcca tatatatatt tgtaatcgtc aacaataaaa acagtttgcc 944
100 Ala
102 acatacatat aaatagtggc taggctcttt tcatccaatt aatatatttt ggcaaagtgt 1004
104 cgatgtttctt atatcatcat atataaatta gcaggtctctt ttcttttttt gta 1057
107 <210> SEQ ID NO: 2
108 <211> LENGTH: 256
109 <212> TYPE: PRT
110 <213> ORGANISM: Arabidopsis thaliana
112 <220> FEATURE:
113 <223> OTHER INFORMATION: APETALA1 (AP1)
115 <400> SEQUENCE: 2
116 Met Gly Arg Gly Arg Val Gln Leu Lys Arg Ile Glu Asn Lys Ile Asn
117 1 5 10 15
118 Arg Gln Val Thr Phe Ser Lys Arg Arg Ala Gly Leu Leu Lys Lys Ala
119 20 25 30
120 His Glu Ile Ser Val Leu Cys Asp Ala Glu Val Ala Leu Val Val Phe
121 35 40 45
122 Ser His Lys Gly Lys Leu Phe Glu Tyr Ser Thr Asp Ser Cys Met Glu
123 50 55 60
124 Lys Ile Leu Glu Arg Tyr Glu Arg Tyr Ser Tyr Ala Glu Arg Gln Leu
125 65 70 75 80
126 Ile Ala Pro Glu Ser Asp Val Asn Thr Asn Trp Ser Met Glu Tyr Asn
127 85 90 95
128 Arg Leu Lys Ala Lys Ile Glu Leu Leu Glu Arg Asn Gln Arg His Tyr
129 100 105 110
130 Leu Gly Glu Asp Leu Gln Ala Met Ser Pro Lys Glu Leu Gln Asn Leu

```

RAW SEQUENCE LISTING

DATE: 10/24/2001

PATENT APPLICATION: US/09/853,450

TIME: 16:01:58

Input Set : A:\-24.app

Output Set: N:\CRF3\10242001\I853450.raw

```

131          115          120          125
132 Glu Gln Gln Leu Asp Thr Ala Leu Lys His Ile Arg Thr Arg Lys Asn
133          130          135          140
134 Gln Leu Met Tyr Glu Ser Ile Asn Glu Leu Gln Lys Lys Glu Lys Ala
135 145          150          155          160
136 Ile Gln Glu Gln Asn Ser Met Leu Ser Lys Gln Ile Lys Glu Arg Glu
137          165          170          175
138 Lys Ile Leu Arg Ala Gln Gln Glu Gln Trp Asp Gln Gln Asn Gln Gly
139          180          185          190
140 His Asn Met Pro Pro Pro Leu Pro Pro Gln Gln His Gln Ile Gln His
141          195          200          205
142 Pro Tyr Met Leu Ser His Gln Pro Ser Pro Phe Leu Asn Met Gly Gly
143          210          215          220
144 Leu Tyr Gln Glu Asp Asp Pro Met Ala Met Arg Arg Asn Asp Leu Glu
145 225          230          235          240
146 Leu Thr Leu Glu Pro Val Tyr Asn Cys Asn Leu Gly Cys Phe Ala Ala
147          245          250          255
150 <210> SEQ ID NO: 3
151 <211> LENGTH: 794
152 <212> TYPE: DNA
153 <213> ORGANISM: Brassica oleracea
155 <220> FEATURE:
156 <221> NAME/KEY: CDS
157 <222> LOCATION: (36)..(794)
158 <223> OTHER INFORMATION: APETALA1 (AP1)
160 <400> SEQUENCE: 3
161 tcttagagga aatagttcct ttaaaaggga taaaa atg gga agg ggt agg gtt      53
162                               Met Gly Arg Gly Arg Val
163                               1           5
165 cag ttg aag agg ata gaa aac aag atc aat aga caa gtg aca ttc tcg      101
166 Gln Leu Lys Arg Ile Glu Asn Lys Ile Asn Arg Gln Val Thr Phe Ser
167          10          15          20
169 aaa aga aga gct ggt ctt atg aag aaa gct cat gag atc tct gtt ctg      149
170 Lys Arg Arg Ala Gly Leu Met Lys Lys Ala His Glu Ile Ser Val Leu
171          25          30          35
173 tgt gat gct gaa gtt gcg ctt gtt gtc ttc tcc cat aag ggg aaa ctc      197
174 Cys Asp Ala Glu Val Ala Leu Val Val Phe Ser His Lys Gly Lys Leu
175          40          45          50
177 ttt gaa tac tcc act gat tct tgt atg gag aag ata ctt gaa cgc tat      245
178 Phe Glu Tyr Ser Thr Asp Ser Cys Met Glu Lys Ile Leu Glu Arg Tyr
179 55          60          65          70
181 gag aga tac tct tac gcc gag aga cag ctt ata gca cct gag tcc gac      293
182 Glu Arg Tyr Ser Tyr Ala Glu Arg Gln Leu Ile Ala Pro Glu Ser Asp
183          75          80          85
185 tcc aat acg aac tgg tcg atg gag tat aat agg ctt aag gct aag att      341
186 Ser Asn Thr Asn Trp Ser Met Glu Tyr Asn Arg Leu Lys Ala Lys Ile
187          90          95          100
189 gag ctt ttg gag aga aac cag agg cac tat ctt ggg gaa gac ttg caa      389
190 Glu Leu Leu Glu Arg Asn Gln Arg His Tyr Leu Gly Glu Asp Leu Gln

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/853,450

DATE: 10/24/2001

TIME: 16:01:58

Input Set : A:\-24.app

Output Set: N:\CRF3\10242001\I853450.raw

```

191          105          110          115
193 gca atg agc cct aag gaa ctc cag aat cta gag caa cag ctt gat act 437
194 Ala Met Ser Pro Lys Glu Leu Gln Asn Leu Glu Gln Gln Leu Asp Thr
195          120          125          130
197 gct ctt aag cac atc cgc tct aga aaa aac caa ctt atg tac gac tcc 485
198 Ala Leu Lys His Ile Arg Ser Arg Lys Asn Gln Leu Met Tyr Asp Ser
199 135          140          145          150
201 atc aat gag ctc caa aga aag gag aaa gcc ata cag gaa caa aac agc 533
202 Ile Asn Glu Leu Gln Arg Lys Glu Lys Ala Ile Gln Glu Gln Asn Ser
203          155          160          165
205 atg ctt tcc aag cag att aag gag agg gaa aac gtt ctt agg gcg caa 581
206 Met Leu Ser Lys Gln Ile Lys Glu Arg Glu Asn Val Leu Arg Ala Gln
207          170          175          180
209 caa gag caa tgg gac gag cag aac cat ggc cat aat atg cct ccg cct 629
210 Gln Glu Gln Trp Asp Glu Gln Asn His Gly His Asn Met Pro Pro Pro
211          185          190          195
213 cca ccc ccg cag cag cat caa atc cag cat cct tac atg ctc tct cat 677
214 Pro Pro Pro Gln Gln His Gln Ile Gln His Pro Tyr Met Leu Ser His
215          200          205          210
217 cag cca tct cct ttt ctc aac atg ggg ggg ctg tat caa gaa gaa gat 725
218 Gln Pro Ser Pro Phe Leu Asn Met Gly Gly Leu Tyr Gln Glu Glu Asp
219 215          220          225          230
221 caa atg gca atg agg agg aac gat ctc gat ctg tct ctt gaa ccc ggt 773
222 Gln Met Ala Met Arg Arg Asn Asp Leu Asp Leu Ser Leu Glu Pro Gly
223          235          240          245
225 tat aac tgc aat ctc ggc tgc 794
226 Tyr Asn Cys Asn Leu Gly Cys
227          250
230 <210> SEQ ID NO: 4
231 <211> LENGTH: 253
232 <212> TYPE: PRT
233 <213> ORGANISM: Brassica oleracea
235 <220> FEATURE:
236 <223> OTHER INFORMATION: APETALA1 (AP1)
238 <400> SEQUENCE: 4
239 Met Gly Arg Gly Arg Val Gln Leu Lys Arg Ile Glu Asn Lys Ile Asn
240 1          5          10          15
241 Arg Gln Val Thr Phe Ser Lys Arg Arg Ala Gly Leu Met Lys Lys Ala
242          20          25          30
243 His Glu Ile Ser Val Leu Cys Asp Ala Glu Val Ala Leu Val Val Phe
244          35          40          45
245 Ser His Lys Gly Lys Leu Phe Glu Tyr Ser Thr Asp Ser Cys Met Glu
246          50          55          60
247 Lys Ile Leu Glu Arg Tyr Glu Arg Tyr Ser Tyr Ala Glu Arg Gln Leu
248 65          70          75          80
249 Ile Ala Pro Glu Ser Asp Ser Asn Thr Asn Trp Ser Met Glu Tyr Asn
250          85          90          95
251 Arg Leu Lys Ala Lys Ile Glu Leu Leu Glu Arg Asn Gln Arg His Tyr
252          100          105          110

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/853,450

DATE: 10/24/2001

TIME: 16:01:58

Input Set : A:\-24.app

Output Set: N:\CRF3\10242001\I853450.raw

```

253 Leu Gly Glu Asp Leu Gln Ala Met Ser Pro Lys Glu Leu Gln Asn Leu
254      115      120      125
255 Glu Gln Gln Leu Asp Thr Ala Leu Lys His Ile Arg Ser Arg Lys Asn
256      130      135      140
257 Gln Leu Met Tyr Asp Ser Ile Asn Glu Leu Gln Arg Lys Glu Lys Ala
258 145      150      155      160
259 Ile Gln Glu Gln Asn Ser Met Leu Ser Lys Gln Ile Lys Glu Arg Glu
260      165      170      175
261 Asn Val Leu Arg Ala Gln Gln Glu Gln Trp Asp Glu Gln Asn His Gly
262      180      185      190
263 His Asn Met Pro Pro Pro Pro Pro Gln Gln His Gln Ile Gln His
264      195      200      205
265 Pro Tyr Met Leu Ser His Gln Pro Ser Pro Phe Leu Asn Met Gly Gly
266      210      215      220
267 Leu Tyr Gln Glu Glu Asp Gln Met Ala Met Arg Arg Asn Asp Leu Asp
268 225      230      235      240
269 Leu Ser Leu Glu Pro Gly Tyr Asn Cys Asn Leu Gly Cys
270      245      250
273 <210> SEQ ID NO: 5
274 <211> LENGTH: 768
275 <212> TYPE: DNA
276 <213> ORGANISM: Brassica oleracea var. botrytis
278 <220> FEATURE:
279 <221> NAME/KEY: CDS
280 <222> LOCATION: (1)..(768)
281 <223> OTHER INFORMATION: APETALA1 (AP1)
283 <400> SEQUENCE: 5
284 atg gga agg ggt agg gtt cag ttg aag agg ata gaa aac aag atc aat 48
285 Met Gly Arg Gly Arg Val Gln Leu Lys Arg Ile Glu Asn Lys Ile Asn
286 1 5 10 15
288 aga caa gtg aca ttc tcg aaa aga aga gct ggt ctt atg aag aaa gct 96
289 Arg Gln Val Thr Phe Ser Lys Arg Arg Ala Gly Leu Met Lys Lys Ala
290 20 25 30
292 cat gag atc tct gtt ctg tgt gat gct gaa gtt gcg ctt gtt gtc ttc 144
293 His Glu Ile Ser Val Leu Cys Asp Ala Glu Val Ala Leu Val Val Phe
294 35 40 45
296 tcc cat aag ggg aaa ctc ttt gaa tac ccc act gat tct tgt atg gag 192
297 Ser His Lys Gly Lys Leu Phe Glu Tyr Pro Thr Asp Ser Cys Met Glu
298 50 55 60
300 gag ata ctt gaa cgc tat gag aga tac tct tac gcc gag aga cag ctt 240
301 Glu Ile Leu Glu Arg Tyr Glu Arg Tyr Ser Tyr Ala Glu Arg Gln Leu
302 65 70 75 80
304 ata gca cct gag tcc gac tcc aat acg aac tgg tcg atg gag tat aat 288
305 Ile Ala Pro Glu Ser Asp Ser Asn Thr Asn Trp Ser Met Glu Tyr Asn
306 85 90 95
308 agg ctt aag gct aag att gag ctt ttg gag aga aac cag agg cac tat 336
309 Arg Leu Lys Ala Lys Ile Glu Leu Leu Glu Arg Asn Gln Arg His Tyr
310 100 105 110
312 ctt ggg gaa gac ttg caa gca atg agc cct aag gaa ctc cag aat cta 384

```

VERIFICATION SUMMARY

DATE: 10/24/2001

PATENT APPLICATION: US/09/853,450

TIME: 16:01:59

Input Set : A:\-24.app

Output Set: N:\CRF3\10242001\I853450.raw

L:618 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:9
L:997 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:15
L:1385 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:19
L:1498 M:300 W: (50) Intentionally skipped Sequence, : Sequence Id (25) SEQUENCE:
L:1504 M:300 W: (50) Intentionally skipped Sequence, : Sequence Id (26) SEQUENCE:
L:1923 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:33
L:1999 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:35
L:2000 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:35
L:2062 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:36
L:2244 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:39
L:2248 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:39
L:2252 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:39
L:2256 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:39
L:2260 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:39